Available online at www.sciencedirect.com
SCIENCE DIRECT®
LINEAR ALGEBRA AND ITS APPLICATIONS
Linear Algebra and its Applications 394 (2005) 91–107
www.elsevier.com/locate/laa

On the accuracy of the ellipsoid norm approximation of the joint spectral radius

Vincent D. Blondel, Yurii Nesterov, Jacques Theys*

Department of Mathematical Engineering, Center for Systems Engineering and Applied Mechanics (CESAME), University of Louvain, B-1348 Louvain-la-Neuve, Belgium

Received 7 November 2003; accepted 26 June 2004

Submitted by H. Schneider

Abstract

The joint spectral radius of a set of matrices is a measure of the maximal asymptotic growth rate that can be obtained by forming long products of matrices taken from the set. This quantity appears in a number of application contexts but is notoriously difficult to compute and to approximate. We introduce in this paper an approximation $\hat{\rho}$ that is based on ellipsoid norms, that can be computed by convex optimization, and that is such that the joint spectral radius belongs to the interval $[\hat{\rho}/\sqrt{n}, \hat{\rho}]$, where n is the dimension of the matrices. We also provide a simple approximation for the special case where the entries of the matrices are non-negative; in this case the approximation is proved to be within a factor at most m (m is the number of matrices) of the exact value.

Keywords: Switched systems; Joint spectral radius; Generalized spectral radius; Stability; Approximation algorithm; Ellipsoid norm

1. Introduction

The *joint spectral radius* of a set $\mathscr{M}$ of matrices is a quantity, introduced by Rota and Strang in the early 60s (see [17]), that measures the maximal asymptotic growth

* Corresponding author.
E-mail addresses: blondel@inma.ucl.ac.be (V.D. Blondel), nesterov@inma.ucl.ac.be (Y. Nesterov), theys@inma.ucl.ac.be (J. Theys).

doi:10.1016/j.laa.2004.06.024

rate that can be obtained by forming long products of matrices taken from $\mathscr{M}$. More formally, the joint spectral radius is defined by:

$$\rho(\mathscr{M}) := \limsup_{k\to\infty} \rho_k(\mathscr{M}),$$

where

$$\rho_k(\mathscr{M}) = \max_{A_1,\ldots,A_k\in\mathscr{M}} \|A_k \cdots A_1\|^{1/k}.$$

The values of $\rho_k(\mathscr{M})$ do in general depend on the chosen norm but one can show that the limit value $\rho(\mathscr{M})$ does not. When the set $\mathscr{M}$ consists of only one matrix, the joint spectral radius coincides with the usual spectral radius, which is equal to the maximum magnitude of the eigenvalues of the matrix. In the previous definition, if we had used the spectral radius instead of the norm, we would have obtained the *generalized spectral radius*:

$$\rho'(\mathscr{M}) = \limsup_{k\to\infty} \rho'_k(\mathscr{M}),$$

where

$$\rho'_k(\mathscr{M}) = \max_{A_1,\ldots,A_k\in\mathscr{M}} \rho(A_k \cdots A_1)^{1/k}.$$

This quantity appears for the first time in [5], where it is also conjectured that in the case of bounded sets of matrices (and in particular for finite sets of matrices), the joint and generalized spectral radii are equal. This conjecture is proved to be correct in [3]. In this paper we shall only be interested in finite sets of matrices and, because of the equality between the joint and generalized spectral radii in this case, we shall sometime refer to this quantity simply as *spectral radius*.

Questions related to the computability of the spectral radius of sets of matrices have been posed in [19] and [12]. The spectral radius can easily be approximated to any desired accuracy. Indeed, the following bounds, proved in [12],

$$\rho'_k(\mathscr{M}) \leqslant \rho(\mathscr{M}) \leqslant \rho_k(\mathscr{M})$$

can be evaluated for increasing values of k and lead to arbitrary close approximations of ρ. These are however expensive calculations. It is proved in [20] that, unless $P = NP$, there is in fact no polynomial-time approximation algorithm for the spectral radius of two matrices. The problem of computing approximations of the joint spectral radius is raised and analyzed in a number of recent contributions. In [14], the exponential number of products that appear in the naive computation of ρ'_k is reduced by avoiding duplicate computation of cyclic permutations; the total number of product to consider remains however exponential. In [7], an algorithm based on the above idea is presented. The algorithm gives arbitrarily small intervals for the joint spectral radius, but no rate of convergence is proved.

In this paper, we provide two easily computable approximations of the spectral radius for finite sets of matrices. The first approximation that we provide, $\hat{\rho}$, is based on the computation of a common quadratic Lyapunov function, or, equivalently, on the computation of an ellipsoid norm. This approximation has the advantage that it can be expressed as a convex optimization problem for which efficient algorithms exist. This first approximation satisfies

$$\frac{1}{\sqrt{n}}\,\hat{\rho} \leqslant \rho \leqslant \hat{\rho},$$

where n is the dimension of the matrices. For the special case of symmetric matrices, triangular matrices, or for sets of matrices that have a solvable Lie algebra, we prove equality between the joint spectral radius and its approximation, $\rho = \hat{\rho}$.

We then prove a result of independent interest: the largest spectral radius of the matrices in the convex hull of $\mathscr{M} = \{A_1, \ldots, A_m\}$ is a lower bound on the spectral radius of $\mathscr{M}$:

$$\max_{0 \leqslant \lambda_i \leqslant 1,\, \sum \lambda_i = 1} \rho\left(\sum_i \lambda_i A_i\right) \leqslant \rho(\mathscr{M}).$$

By using this inequality, we prove a simple bound for the spectral radius of sets of matrices that have *non-negative entries*. The spectral radius of the matrix S whose entries are the componentwise maximum of the entries of the matrices in $\mathscr{M}$ satisfies

$$\frac{\rho(S)}{m} \leqslant \rho(\mathscr{M}) \leqslant \rho(S),$$

where m is the number of matrices in the set. In this expression, $\mathscr{M}$ is a set of matrices, whereas S is a single matrix.

The paper is organized as follows. In the next section, we recall some useful results regarding the joint spectral radius. In Section 3, we define the spectral radius approximation based on ellipsoid norms. In Section 4, we describe situations for which this approximation is exact, and situations for which it is not. In Section 5, we prove the inequality $\hat{\rho}(\mathscr{M})/\sqrt{n} \leqslant \rho(\mathscr{M})$ by using a geometrical property of ellipsoids known as John's ellipsoid theorem. Finally, in Section 6 we provide an under-approximation of the joint spectral radius based on the spectral radius of all convex combinations of the matrices in the set $\mathscr{M}$ and use this result to prove an approximation for sets of non-negative matrices.

2. Useful properties of the joint spectral radius

The joint spectral radius can be defined by an extremal norm property. The following theorem is a result from Barabanov in [2]. It is stated here as a founding result, and only to introduce Theorem 3. An extensive discussion of extremal norms by Wirth can be found in [21], where the statement of Barabanov's theorem also appears. Before proceeding, let us recall the notion of *irreducibility* of a set of matrices: a set $\mathscr{M}$ of matrices is *irreducible* if only the trivial subspaces $\{0\}$ and $\mathbb{R}^n$ are invariant under all matrices $A_i \in \mathscr{M}$.

Theorem 1 (Barabanov). *Let $\rho(\mathscr{M})$ be the joint spectral radius of the finite set of matrices $\mathscr{M} = \{A_1, \ldots, A_m\}$. If $\mathscr{M}$ is irreducible, then there exists a vector norm $\|\cdot\|_*$ for which*:

(1) $\|A_i x\|_* \leqslant \rho(\mathscr{M})\ \|x\|_*$, $\forall x$ *and* $\forall A_i \in \mathscr{M}$;
(2) $\forall x, \exists A_i \in \mathscr{M} : \|A_i x\|_* = \rho(\mathscr{M})\ \|x\|_*$.

Notice that, considering the matrix norm induced by this vector norm, and also denoting it by $\|\cdot\|_*$, we have

$$\|A_i\|_* \leqslant \rho(\mathscr{M}), \quad \forall i. \tag{1}$$

We can now widen the scope of the previous theorem, by removing the irreducibility hypothesis. To do so, we make use of a result from Kozyakin in [11]. We first recall his own definition of stability.

Definition 2. The system $x_k = A_{i(k)}x_{k-1}$ is said to be *absolutely exponentially stable* with respect to a matrix class $\mathscr{U}$ if for a $c = c(\mathscr{U})$ and a $q = q(\mathscr{U}) < 1$, the following inequality holds for any sequence $(A_i) \in \mathscr{U}$

$$\|A_{i(k)}A_{i(k-1)} \cdots A_{i(1)}\ x\| \leqslant cq^k\|x\|\ .$$

Theorem 3 (Kozyakin). *The system $x_k = A_{i(k)}x_{k-1}$ with $A_{i(k)} \in \{A_1, A_2, \ldots, A_m\}$ is absolutely exponentially stable iff there exists a norm $\|\cdot\|_*$ and a scalar $q < 1$ such that*

$$\|A_1\|_*, \|A_2\|_*, \ldots, \|A_m\|_* \leqslant q.$$

It is worth noting that Theorem 3 completes Theorem 1. Both can be combined in the following statement.

Proposition 4. *For any finite set of matrices $\mathscr{M} = \{A_1, \ldots, A_m\}$, for any $\varepsilon > 0$, there exists a matrix norm $\|\cdot\|_*$ such that*

$$\|A_1\|_*, \|A_2\|_*, \ldots, \|A_m\|_* \leqslant \rho(\mathscr{M}) + \varepsilon.$$

If the set is irreducible, then there exists a matrix norm $\|\cdot\|_*$ such that

$$\|A_1\|_*, \|A_2\|_*, \ldots, \|A_m\|_* \leqslant \rho(\mathscr{M}).$$

Proof. Consider a finite set $\mathscr{M}$ of matrices of joint spectral radius ρ^*. Let us normalize it with any quantity larger than ρ^*:

$$\forall r > \rho^*, \quad \rho(\mathscr{M}/r) < 1.$$

One can then apply Theorem 3: there exists a $q < 1$ and a norm $\|\cdot\|_*$ such that $\left\|\frac{A_i}{r}\right\|_* \leqslant q, \forall i$ (i.e. $\max_i \left\|\frac{A_i}{r}\right\|_* \leqslant 1, \forall x$). By linearity of the norm and the joint spectral radius, this allows us to deduce that

$$\forall r > \rho(\mathscr{M}), \quad \exists \|\cdot\|_* : \max_i \|A_i\|_* \leqslant r.$$

One can replace r by $\rho(\mathscr{M}) + \varepsilon$ to obtain the exact formulation of the proposition.

The second statement is simply Barabanov's theorem. □

Remark 5. It is important to mention that, in the proof of Theorem 3 (see [11]), it is established that absolute exponential stability implies the existence of a *vector* norm $\|\cdot\|_*$ such that

$$\|A_1 x\|_*, \|A_2 x\|_*, \ldots, \|A_m x\|_* \leqslant q\|x\|_*, \quad \forall x .$$

So, we could have expressed the inequalities in Theorem 3 and Proposition 4 using vector norms, as here above. Further on, we will make use of such a vector norm version of Proposition 4.

The joint spectral radius can thus be seen as the infimum over all possible matrix norms of the largest norm of the matrices in the set. A norm achieving this infimum is said to be *extremal* for the set (not every set of matrices possesses an extremal norm, see [21] for a discussion of this issue). In [11], Kozyakin describes the theoretical construction of such an extremal norm. This construction is not explicit and partly relies on the apriori knowledge of the numerical value of the joint spectral radius. One can of course not hope enumerating all possible matrix norms for computing the joint spectral radius, but we can enumerate particular sets of norms. This is what we do in the next section.

3. The ellipsoid norm approximation: definition

Following the previous discussion, our first approximation of the joint spectral radius is obtained by finding, among all ellipsoid norms $\|\cdot\|_P$, one that minimizes $\max_i \|A_i\|_P$.

Let us briefly recall the definition of the ellipsoid (or quadratic) norms. Let P be a positive definite matrix;[1] the vector P-norm is defined as $\|x\|_P = \sqrt{x^{\mathrm{T}} P x}$. Associated to this vector norm, there is an induced matrix norm:

$$||A_i||_P = \max_x \frac{\|A_i x\|_P}{\|x\|_P} = \max_x \frac{\sqrt{x^{\mathrm{T}} A_i^{\mathrm{T}} P A_i x}}{\sqrt{x^{\mathrm{T}} P x}}. \tag{2}$$

[1] Positive definiteness is denoted $\succ 0$ and positive semi-definiteness is denoted $\succeq 0$.

Note that we use the notation $\|\cdot\|_P$ for both the vector and matrix norms, as there is no ambiguity regarding which one is applied. Let us now define the *ellipsoid norm approximation* of the joint spectral radius by:

$$\hat{\rho}(\mathscr{M}) = \inf_{P \succ 0} \max_{A_i \in \mathscr{M}} \|A_i\|_P.$$

The minimum on all ellipsoid norms cannot be lower than the minimum on all possible norms and so it immediately follows from the discussion after Proposition 4 that $\rho(\mathscr{M}) \leqslant \hat{\rho}(\mathscr{M})$. The ellipsoid norm approximation can be computed as follows. Notice first that Definition (2) implies that

$$\begin{aligned}
&\forall x, \quad \sqrt{x^{\mathrm{T}} A_i^{\mathrm{T}} P A_i x} \leqslant \|A_i\|_P \sqrt{x^{\mathrm{T}} P x} \\
&\forall x, \quad x^{\mathrm{T}} A^{\mathrm{T}} P A x \leqslant \|A\|_P^2 \, x^{\mathrm{T}} P x \\
&\forall x, \quad x^{\mathrm{T}} (A_i^{\mathrm{T}} P A_i - \|A_i\|_P^2 P) x \leqslant 0 \\
&A_i^{\mathrm{T}} P A_i - \|A_i\|_P^2 P \preceq 0.
\end{aligned}$$

One can therefore think of $\|A_i\|_P$ as the smallest scalar value γ for which $A_i^{\mathrm{T}} P A_i \preceq \gamma^2 P$ for some $P \succ 0$. The ellipsoid norm approximation of a set $\mathscr{M} = \{A_1, \ldots, A_m\}$ is thus equal to the smallest scalar γ for which there is a solution $P \succ 0$ to $A_i^{\mathrm{T}} P A_i \preceq \gamma^2 P$, $\forall i$. This problem can be solved efficiently by convex optimization applied to linear matrix inequalities (LMI); see, e.g. [15].

A natural question to ask is how good this approximation is in the general case. In the next section, we describe situations for which the approximation is equal to the joint spectral radius, and we provide an example for which the approximation is larger than the joint spectral radius.

4. The ellipsoid norm approximation: special cases

We prove in this section that the joint spectral radius and the ellipsoid norm approximation are equal (and are equal to the maximal spectral radius of the matrices in the set) in the following situations: all matrices are symmetric, all matrices are triangular or, more generally, the Lie algebra associated to the matrices is solvable. We close the section with an example for which the spectral radius and its approximation are different. We start with the case of symmetric matrices:

Proposition 6. *For a set of symmetric matrices, the spectral radius and its ellipsoid norm approximation are equal. Their value is the largest spectral radius of the matrices in the set.*

Proof. Using the identity I as matrix P, we get $\|A_i\|_I = \inf\{\gamma : A_i^2 \preceq \gamma^2 I\}$ and $\|A_i\|_I = \rho(A_i)$. Knowing that $\rho(A_i) \leqslant \inf_{P \succ 0} \|A_i\|_P$, we have actually strict equality $\rho(A_i) = \inf_{P\succ 0} \|A_i\|_P$. Finally, we have that $\hat{\rho}(\mathscr{M}) = \inf_{P\succ 0} \max_i \|A_i\|_P = \max_i \rho(A_i)$. $\square$

In order to derive our result for triangular matrices, we first establish a discrete-time analog to a continuous-time result of [13] on the existence of a common quadratic Lyapunov function for switched linear systems.

Lemma 7. *Let $\mathscr{M}$ be the set $\{A_1, \ldots, A_m\}$ and consider the discrete-time switched linear system*

$$x_{k+1} = A_{i_k} x_k, \quad A_{i_k} \in \mathscr{M}.$$

If the system is stable and the matrices are upper-triangular, then there exists a common quadratic Lyapunov function in the form of a diagonal matrix.

Proof. Let $\{A_i, \ldots, A_m\}$ be a set of upper-triangular (possibly complex) matrices and P the candidate Lyapunov function (diagonal, real):

$$A_i = \begin{pmatrix} a_{11}^i & a_{12}^i & \cdots & a_{1n}^i \\ 0 & a_{22}^i & \cdots & a_{2n}^i \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^i \end{pmatrix}, \quad P = \begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{pmatrix}, \quad p_k > 0, \forall k.$$

For P to be a Lyapunov function of $x_{k+1} = A_i x_k$ (fixed A_i), the following relation has to hold:

$$P - A_i^* P A_i \succ 0.$$

Developing $P - A_i^* P A_i$, we get:

$$\begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{pmatrix} - \begin{pmatrix} {a_{11}^i}^* & 0 & \cdots & 0 \\ {a_{12}^i}^* & {a_{22}^i}^* & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ {a_{1n}^i}^* & {a_{2n}^i}^* & \cdots & {a_{nn}^i}^* \end{pmatrix} \begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{pmatrix}$$

$$\times \begin{pmatrix} a_{11}^i & a_{12}^i & \cdots & a_{1n}^i \\ 0 & a_{22}^i & \cdots & a_{2n}^i \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^i \end{pmatrix} = \begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{pmatrix}$$

$$
-\begin{pmatrix} {a^i_{11}}^* p_1 & 0 & \cdots & 0 \\ {a^i_{12}}^* p_1 & {a^i_{22}}^* p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ {a^i_{1n}}^* p_1 & {a^i_{2n}}^* p_2 & \cdots & {a^i_{nn}}^* p_n \end{pmatrix} \begin{pmatrix} a^i_{11} & a^i_{12} & \cdots & a^i_{1n} \\ 0 & a^i_{22} & \cdots & a^i_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a^i_{nn} \end{pmatrix}
$$

which yields

$$
\begin{pmatrix} (1-|a^i_{11}|^2)p_1 & -{a^i_{11}}^* a^i_{12} p_1 & \cdots \\ -a^i_{11}{a^i_{12}}^* p_1 & -|a^i_{12}|^2 p_1 + (1-|a^i_{22}|^2)p_2 & \cdots \\ \vdots & \vdots & \ddots \\ -a^i_{11}{a^i_{1n}}^* p_1 & -a^i_{12}{a^i_{1n}}^* p_1 - a^i_{22}{a^i_{2n}}^* p_2 & \cdots \end{pmatrix}. \tag{3}
$$

The first thing to note is that this matrix is Hermitian, and so its leading principal minors are real (see [8]).

As A_i is assumed to be stable, $a^i_{jj} < 1, \forall j$. The first diagonal element in (3) is therefore positive, for any value of p_1. Let it be chosen as 1. Moreover, the value of p_2 can be chosen in such a way that the (2×2) leading principal minor is positive. Indeed, p_2 only appears in its last diagonal element, and its coefficient $(1 - |a^i_{22}|^2)$ is positive, as $a^i_{22} < 1$. So, taking p_2 such that

$$
\begin{vmatrix} (1-|a^i_{11}|^2) & -{a^i_{11}}^* a^i_{12} \\ -a^i_{11}{a^i_{12}}^* & -|a^i_{12}|^2 + (1-|a^i_{22}|^2)p_2 \end{vmatrix} > 0
$$

is possible, and simple developments give the following condition:

$$
p_2 > \frac{1}{(1-|a^i_{22}|^2)} \left[\frac{(|a^i_{11}||a^i_{12}|)^2}{(1-|a^i_{11}|^2)} + |a^i_{12}|^2 \right].
$$

We can define in this way a p_2 that satisfies this for all matrices A_i of the set by choosing

$$
p_2 > \max_i \frac{1}{(1-|a^i_{22}|^2)} \left[\frac{(|a^i_{11}||a^i_{12}|)^2}{(1-|a^i_{11}|^2)} + |a^i_{12}|^2 \right].
$$

The same argument shows that we can successively choose the values of $p_3, \ldots, p_n$ in a way such that the leading principal minors of (3) are all positive, and this for any matrix A_i of the set. Indeed, let the leading principal minor of order k be > 0. Then, the leading principal minor of order $k + 1$ can be made > 0 too, because p_{k+1} only appears in its last diagonal term, with a strictly positive coefficient. So, taking p_{k+1} large enough is sufficient. The finiteness of the elements of A_i guarantees us that such a value p_{k+1} exists and is finite.

A Hermitian matrix H is positive definite if and only if all its leading principal minors are positive ([8]), and so we can deduce that the Hermitian matrix appearing in (3) is indeed positive definite, for any i. So, the P matrix built in this way is a common quadratic Lyapunov function for the set $\mathscr{M}$. $\square$

Corollary 8. *For a set $\mathscr{M}$ of triangular matrices A_i, the spectral radius of $\mathscr{M}$ is equal to* $\max_i \rho(A_i)$ *and also to its ellipsoid norm approximation.*

Proof. From Lemma 7, it follows that, for a set of stable upper-triangular matrices A_i, there exists a positive definite P_* such that $\|A_i\|_{P_*} < 1, \forall i$. This is equivalent to expressing

$$\max_i \rho(A_i) < 1 \Rightarrow \exists P_* \succ 0 : \max_i \|A_i\|_{P_*} < 1.$$

By linearity, this implies that $\max_i \rho(A_i) \geqslant \max_i \|A_i\|_{P_*}$. Indeed, let us pose $\max_i \rho(A_i) = r$, so that

$$\forall y > r, \quad \max_i \rho\left(\frac{A_i}{y}\right) < 1.$$

This implies that $\forall y > r, \exists P_* : \max_i \frac{\|A_i\|_{P_*}}{y} < 1$ or again,

$$\forall y > \max_i \rho(A_i), \exists P_* : \max_i \|A_i\|_{P_*} < y.$$

So, $\max_i \|A_i\|_{P_*}$ is arbitrarily close (from above) to $\max_i \rho(A_i)$ and the announced inequality $\max_i \rho(A_i) \geqslant \max_i \|A_i\|_{P_*}$ holds.

On the other hand, we know that the joint spectral radius is greater or equal to the largest spectral radius of the matrices in the set, that is $\rho(\mathscr{M}) \geqslant \max_i \rho(A_i)$. So, summing up, we have

$$\rho(\mathscr{M}) \geqslant \max_i \rho(A_i) \geqslant \max_i \|A_i\|_P \geqslant \hat{\rho}(\mathscr{M}).$$

As $\hat{\rho}$ is an over-approximation of $\rho(\mathscr{M})$, the four members of this inequality are equal. This yields $\rho(\mathscr{M}) = \max_i \rho(A_i)$ and $\hat{\rho}(\mathscr{M}) = \rho(\mathscr{M})$. $\square$

We now generalize the previous result to a more general class of sets of matrices. This development is very similar to the one presented in [13]. Let us recall the following notations and definitions. The *commutator* of the matrices A, B is the matrix given by $[A, B] = AB - BA$. The Lie algebra $\{A_0, A_1\}_{LA}$ is the linear span of the set of all possible combinations of commutators

$$\{A_0, A_1, [A_0, A_1], [A_0, [A_0, A_1]], [A_1, [A_0, A_1]], \ldots\}.$$

The *commutator series* of the Lie algebra $\{A_0, A_1\}_{LA}$ is the sequence of subalgebras recursively defined by $\mathfrak{g}^0 = \{A_0, A_1\}_{LA}$, and $\mathfrak{g}^{k+1} = [\mathfrak{g}^k, \mathfrak{g}^k]$. We have

$$\mathfrak{g}^0 \supseteq \mathfrak{g}^1 \supseteq \mathfrak{g}^2 \supseteq \cdots,$$

and if $\mathfrak{g}^k = \mathfrak{g}^{k+1}$ then all subsequent subalgebras are also equal to $\mathfrak{g}^k$. A Lie Algebra is *solvable* if its commutator series $\mathfrak{g}^k$ vanishes for some k.

An often used example of solvable Lie algebra is the vector space of upper-triangular matrices. It is easy to check that the sequence of subalgebras $\mathfrak{g}^k$ is the set of upper-triangular matrices whose elements on the diagonal at distance less than k from the main diagonal are all zero:

$$\mathfrak{g}^0 = \left\{ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \quad a_{ij} \in \mathbb{C}^{n\times n} \right\},$$

$$\mathfrak{g}^1 = \left\{ \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad a_{ij} \in \mathbb{C}^{n\times n} \right\}, \ldots$$

We make use of the following result (cited in [13], referring to [16]):

Lemma 9. *Let $\mathfrak{g}$ be a solvable Lie algebra over an algebraically closed field, and let ρ be a representation of $\mathfrak{g}$ on a vector space V of finite dimension n. Then there exists a basis $\{v_1, \ldots, v_n\}$ of V such that for each $X \in \mathfrak{g}$ the matrix of $\rho(X)$ in that basis takes the upper-triangular form*

$$\begin{pmatrix} \lambda_1(X) & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n(X) \end{pmatrix},$$

where the $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of the matrix $\rho(X)$.

Theorem 10. *Let $\mathscr{M} = \{A_1, \ldots, A_m\}$ and consider the switched linear system*

$$x_{k+1} = A_{i_k} x_k, \quad A_{i_k} \in \mathscr{M}.$$

If all matrices in $\mathscr{M}$ have a spectral radius less than 1 and the Lie algebra associated to $\mathscr{M}$ is solvable, then the system has a common quadratic Lyapunov function.

Proof. So, if $\{A_i : A_i \in \mathscr{M}\}_{LA}$ is solvable, then there exists a (possibly complex) invertible matrix T such that

$$A_i = T^{-1}\tilde{A}_i T, \quad \text{with } \tilde{A}_i \text{ upper-triangular}, \ \forall i.$$

This introduction of complex values does not change the main argument.

Lemma 7 shows that there exists a real common quadratic Lyapunov function $\tilde{P}$ in diagonal form for such a set of matrices $\tilde{\mathscr{M}} = \{\tilde{A}_1, \ldots, \tilde{A}_n\}$. From this $\tilde{P}$, we can deduce the form of the corresponding P for the non-upper-triangular set $\mathscr{M} = \{A_1, \ldots, A_n\}$. To this aim, we rewrite the condition for the existence of a common quadratic Lyapunov function.

$$\tilde{A}_i^* \tilde{P} \tilde{A}_i - \tilde{P} \prec 0,$$

$$(TA_iT^{-1})^*\tilde{P}TA_iT^{-1} - \tilde{P} \prec 0,$$
$$T^{*-1}A_i^*T^*\tilde{P}TA_iT^{-1} - \tilde{P} \prec 0,$$
$$A_i^*(T^*\tilde{P}T)A_i - (T^*\tilde{P}T) \prec 0.$$

And we get $P = T^*\tilde{P}T$. As $\tilde{P}$ is positive definite, so is P. Moreover, $\tilde{P}$ being diagonal, $T^*\tilde{P}T$ is actually Hermitian, but is not guaranteed to be real. Let us then denote

$$-R := A_i^*(T^*\tilde{P}T)A_i - (T^*\tilde{P}T)$$

where R is, by construction, Hermitian positive definite. We can write, by separating the real and imaginary parts,

$$P = \mathbb{R}(P) + i\mathbb{I}(P) \quad \text{and} \quad R = \mathbb{R}(R) + i\mathbb{I}(R).$$

As P and R are Hermitian, $\mathbb{R}(P)$, $\mathbb{R}(R)$ are symmetric positive definite and $\mathbb{I}(P)$, $\mathbb{I}(R)$ are skew-symmetric. We can rewrite

$$A_i^*(\mathbb{R}(P) + i\mathbb{I}(P))A_i - (\mathbb{R}(P) + i\mathbb{I}(P)) = -(\mathbb{R}(R) + i\mathbb{I}(R))$$

and taking the real part,

$$A_i^*\mathbb{R}(P)A_i - \mathbb{R}(P) = -\mathbb{R}(R).$$

As a consequence, $\mathbb{R}(P)$ is a real common quadratic Lyapunov function for the solvable Lie algebra $\{A_i : A_i \in \mathscr{M}\}_{LA}$. $\square$

Corollary 11. *If the Lie algebra associated to the set $\mathscr{M} = \{A_1, \ldots, A_m\}$ is solvable, then the joint spectral radius of $\mathscr{M}$ is equal to* $\max_i \rho(A_i)$ *and also to its ellipsoid norm approximation.*

Proof. Indeed, Theorem 10 allows us to deduce that $\rho(\mathscr{M}) < 1 \Rightarrow \hat{\rho}(\mathscr{M}) < 1$, which yields $\rho(\mathscr{M}) \geqslant \hat{\rho}(\mathscr{M})$, allowing to deduce the strict equality, thanks to the already known $\rho(\mathscr{M}) \leqslant \hat{\rho}(\mathscr{M})$. Here again, we already know that $\rho(\mathscr{M}) \geqslant \max_i \rho(A_i)$, and Theorem 10 teaches us that $\max_i \rho(A_i) < 1 \Rightarrow \rho(\mathscr{M}) < 1$, so $\max_i \rho(A_i) \geqslant \rho(\mathscr{M})$. And we deduce $\rho(\mathscr{M}) = \max_i \rho(A_i)$. $\square$

We have equality between the joint spectral radius and its ellipsoid norm approximation when the Lie algebra is solvable. One could wonder whether the solvability of the Lie algebra is necessary for this equality to hold. This is not the case. In order to exhibit a counter-example, we first prove a property of independent interest.

Proposition 12. *The joint spectral radius of $\{A, A^{\mathrm{T}}\}$ is equal to its ellipsoid norm approximation and to the largest singular value of A, denoted $\sigma(A)$.*

Proof. We use the inequalities

$$\rho(A, A^{\mathrm{T}}) \leqslant \hat{\rho}(A, A^{\mathrm{T}}) \leqslant \sigma(A),$$

which can be deduced by using $P = I$ in the definition of $\hat{\rho}$, so that we get $A^{\mathrm{T}}IA \leqslant \gamma^2 I$, which holds for $\gamma \geqslant \sigma(A)$. And we also have

$$\rho(A, A^{\mathrm{T}}) \geqslant \rho(AA^{\mathrm{T}})^{1/2} = \sigma(A)$$

because AA^{T} is among the products of length 2, taken from the set $\{A, A^{\mathrm{T}}\}$. This eventually yields $\rho(A, A^{\mathrm{T}}) = \hat{\rho}(A, A^{\mathrm{T}}) = \sigma(A)$. $\square$

As an illustration, consider the matrices

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad A^{\mathrm{T}} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

It is easy to check that the Lie algebra associated to these matrices is not solvable. On the other hand it follows from the above proposition that for this pair of matrices the spectral radius and its ellipsoid norm approximation are equal (and are equal to $\sigma(A) = \frac{1+\sqrt{5}}{2} \simeq 1.618$). This provides the counter-example we mentioned above.

We close this section with a numerical example of two matrices for which we do not have equality between the spectral radius and its ellipsoid norm approximation. Let us consider the following matrices (inspired by [6]):

$$A_1 = \begin{pmatrix} 1 & 2\,a_1 \\ -2/a_1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 2\,a_2 \\ -2/a_2 & 1 \end{pmatrix}.$$

Assuming that $a_2 \geqslant a_1 \geqslant 1$, extensive calculations that are not reproduced here (see the Technical Report [18] for more details) show that the approximation $\hat{\rho}$ is such that

$$\hat{\rho}(A_1, A_2) \geqslant \sqrt{1 + 4\,a_2/a_1}.$$

For $a_1 = 1, a_2 = 2$, the joint spectral radius can be shown to be strictly less than 2.8584 by using an exhaustive calculation of all the products of 5 matrices. This is strictly less than $\sqrt{1 + 4 \times 2/1} = 3$, so here $\rho < \hat{\rho}$. The gap between the spectral radius and its approximation can be seen on figure 1 for $a_1 = 1$ and varying a_2.

5. Guaranteed precision of the ellipsoid norm approximation

The ellipsoid norm approximation of the joint spectral radius can be shown to be of guaranteed precision. The argument is simple: Let ρ be the spectral radius of the set $\{A_1, \ldots, A_m\}$. We know by Proposition 4 that, $\forall r > \rho(\mathscr{M})$, there exists a vector norm $\|\cdot\|_*$ for which $\|A_i x\|_* \leqslant r\|x\|_*$ for all x and i. The level curves of this norm define closed convex set that can be approximated by ellipsoids. The quality of these approximations can be measured and provides a guaranteed precision for the approximation.

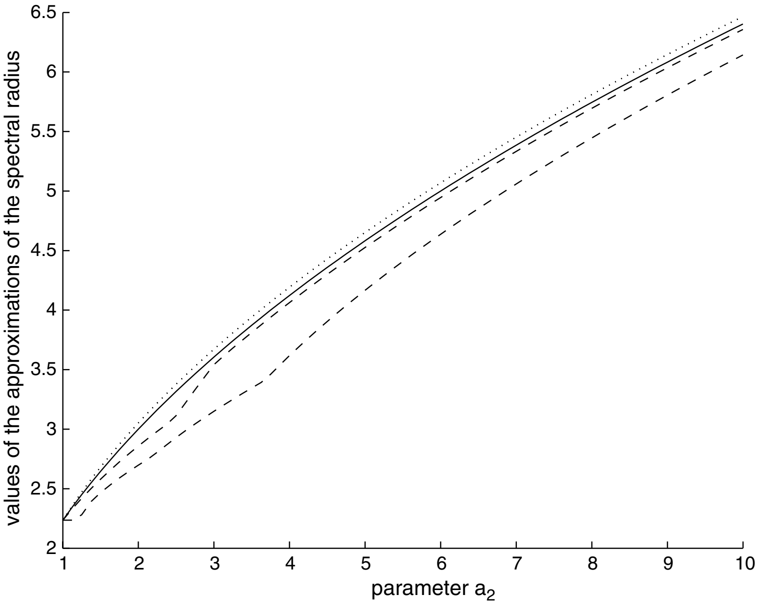

Fig. 1. The spectral radius and its ellipsoid norm approximation as functions of the real parameter a_2 (a_1 is fixed to 1). The two lowest curves (dashed) represent upper and lower bounds on the exact value of the spectral radius (computed with words of length 6). The middle curve (solid) represents $\sqrt{1+4a_2}$. The two highest curves (dotted) represent upper and lower bounds on the approximation.

We start by describing the quality of the best possible ellipsoids (the result below is known as John's theorem; it is stated in [9], referring to [10]).

Theorem 13. *Let $K \subset \mathbb{R}^n$ be a compact convex set with nonempty interior. Then there is an ellipsoid E with center c such that the inclusions $E \subseteq K \subseteq n(E-c)$ hold. If K is symmetric about the origin $(K=-K)$, the constant n can be changed into $\sqrt{n}$.*

Knowing this, we can now prove:

Theorem 14. *Let ρ be the joint spectral radius of a finite set of matrices of dimension n. Let $\hat{\rho}$ be the ellipsoid norm approximation of the joint spectral radius. Then $\hat{\rho}(\mathscr{M})/\sqrt{n} \leqslant \rho(\mathscr{M}) \leqslant \hat{\rho}(\mathscr{M})$.*

Proof. The norm mentioned above is symmetric about the center, as $\|x\|_* = \|-x\|_*, \forall x$. So, Theorem 13 guarantees us that, whatever the norm $\|\cdot\|_*$ is, there exists a quadratic norm $\|x\|_P = x^{\mathrm{T}}Px$ (of which level curves are ellipsoids) such that

$$\|x\|_P \leqslant \|x\|_* \leqslant \sqrt{n}\|x\|_P.$$

As, from Proposition 4 and Remark 5, $\forall q > \rho(\mathscr{M})$, the norm $\|\cdot\|_*$ satisfies $\|A_i x\|_* \leqslant q\|x\|_*, \forall x, \forall i$, we can now write

$$\begin{aligned}
&\forall x, \forall i,\ \|A_i x\|_P \leqslant \|A_i x\|_* \leqslant q\|x\|_* \leqslant q\|x\|_P\sqrt{n},\\
&\forall x, \forall i,\ \ \|A_i x\|_P \leqslant q\|x\|_P\sqrt{n},\\
&\forall x, \forall i,\ \ x^{\mathrm{T}} A_i^{\mathrm{T}} P A_i x \leqslant q^2 n\, x^{\mathrm{T}} P x,\\
&\forall i,\ A_i^{\mathrm{T}} P A_i - q^2 n P \preceq 0.
\end{aligned}$$

Thus, the approximation $\hat{\rho}$ defined by $\hat{\rho}(\mathscr{M}) = \inf_{P \succ 0} \max_{A_i \in \mathscr{M}} \|A_i\|_P$ is $\leqslant q\sqrt{n}$, and this $\forall q > \rho(\mathscr{M})$. So, at worst, the approximation will result in the value $\rho(\mathscr{M})\sqrt{n}$.

On the other hand, as said in Section 3, $\hat{\rho}(\mathscr{M})$ is an over-approximation of $\rho(\mathscr{M})$. This settles the second inequality.

Summing up, this gives:

$$\rho(\mathscr{M}) \leqslant \hat{\rho}(\mathscr{M}) \leqslant \rho(\mathscr{M})\sqrt{n}. \qquad \square$$

6. Matrices with non-negative entries

In this section, we introduce an approximation of the joint spectral radius for matrices with non-negative entries. We first provide a result for general matrices that is of independent interest.

Proposition 15. *Let $\mathscr{M} = \{A_1, \ldots, A_m\}$. Then*

$$\max_{\sum_{i=0}^m \alpha_i = 1,\ \alpha_i \geqslant 0} \rho\left(\sum \alpha_i A_i\right) \leqslant \rho(\mathscr{M}).$$

Proof. We have, using the inequality $\rho(.) \leqslant \|\cdot\|$ (for any valid matrix norm $\|\cdot\|$) and the subadditivity of the norm,

$$\begin{aligned}
\tilde{\rho}(\mathscr{M}) := \rho\left(\sum_i \alpha_i A_i\right) &\leqslant \left\|\sum_i \alpha_i A_i\right\|\\
&\leqslant \sum_i \|\alpha_i A_i\| = \sum_i \alpha_i \|A_i\|\\
&\leqslant \max_i(\|A_i\|),\ \text{ as } \sum_i \alpha_i = 1.
\end{aligned}$$

Now, if we suppose that $\rho(\mathscr{M}) < 1$, we know from Theorem 3 that there exists a norm $\|\cdot\|_*$ such that $\forall i$, $\|A_i\|_* < 1$. We can then deduce from the previous inequality that $\tilde{\rho}(\mathscr{M}) < 1$. So,

$$\rho(\mathscr{M}) < 1 \Rightarrow \tilde{\rho}(\mathscr{M}) < 1,$$

and we deduce, using linearity, that $\tilde{\rho}(\mathscr{M}) \leqslant \rho(\mathscr{M})$. □

We are now ready to prove:

Theorem 16. *Let $\mathscr{M} = \{A_1, \ldots, A_m\}$ be a set of matrices with non-negative entries and define $S_{ij} = \max_{1\leqslant k\leqslant m}(A_k)_{ij}$. We have*

$$\frac{\rho(S)}{m} \leqslant \rho(\mathscr{M}) \leqslant \rho(S). \tag{4}$$

Proof. For the first inequality of (4) we use non-negativity of the matrices, Proposition 15, and the fact that matrices with non-negative entries M_1 and M_2 for which the componentwise inequalities $M_2 \geqslant M_1$ are satisfied are such that $\rho(M_2) \geqslant \rho(M_1)$. From this it follows that

$$S \leqslant \sum_{k=1}^{m} A_k \Rightarrow \rho(S) \leqslant \rho\left(\sum_{k=1}^{m} A_k\right) = m\rho\left(\frac{\sum_{k=1}^{m} A_k}{m}\right) \leqslant m\rho(\mathscr{M}).$$

To prove the second inequality of (4), we may note that, as the elements are non-negative, for any sequence $\omega = (\omega_1, \ldots, \omega_k)$ of k indices, the product $A_\omega = A_{\omega_1} \ldots A_{\omega_k}$ satisfies $(S^k)_{ij} \geqslant (A_\omega)_{ij}$. As above, this allows us to deduce, $\forall\omega : |\omega| = k$,

$$S^k \geqslant A_\omega \Rightarrow \limsup_{k\to\infty} \|S^k\|^{1/k} \geqslant \limsup_{k\to\infty}\left(\max_{|\omega|=k} \|A_\omega\|^{1/k}\right) \Rightarrow \rho(S) \geqslant \rho(\mathscr{M}).$$

□

For any set cardinality m, the equality $\rho(S)/m = \rho(\mathscr{M})$ is attained for particular matrices. For $m = 2$ consider the following pair:

$$\left\{A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\right\}$$

for which $\rho(S) = 2$. On the other hand, we see that $A^2 = A$, $B^2 = B$, $AB = B$ and $BA = A$. Any product generated by $\{A, B\}$ is either A or B, so $\rho(A, B) = 1$ and we indeed have $\hat{\rho}/2 = \rho$. A similar construction is immediate for the cases $m \geqslant 3$.

7. Conclusion

We introduce in this paper a polynomial-time approximation of the joint spectral radius that is easy to compute and that is guaranteed to be within a factor $\sqrt{n}$ of the exact value, where n is the dimension of the matrices. We describe particular classes of matrices for which our approximation is equal to the joint spectral radius. The problem of characterizing exactly the sets of matrices for which equality holds is a question that remains open. We also provide an easy way of approximating the joint spectral radius of matrices with non-negative entries, and show that this approximation is within a factor at most m of the exact value, where m is the number of matrices in the set. This last result does not depend on the size of the matrices. The question remains open to find better approximations at a reasonable computational cost. In particular, both approximations presented in this paper have relative errors that increase with the size or number of the matrices. It is yet unclear if a polynomial time approximation is possible that gives a fixed guaranteed relative error.

Acknowledgments

JT holds a FNRS fellowship (Belgian National Research Fund).

This paper presents research results of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its author(s).

We would like to thank Professor Alexander Megretski for suggesting the proof of Theorem 4 in a discussion at the Mittag-Leffler Institute (Stockholm, Sweden). Thanks are also due to Professor Alexander Vladimirov for numerous discussions during his visits at CESAME in Belgium.

After this paper was submitted and reviewed, the attention of the authors was drawn to the Ref. [1], which contains alternative proofs for results analogous to some of those presented in this contribution.

After this paper was submitted, two of the authors have pursued the present analysis and have derived approximations of arbitrary accuracy, thus improving the quality of the approximation of the joint spectral radius. These results will be subject to a future publication; see [4].

References

[1] T. Ando, M.-h. Shih, Simultaneous contractibility, SIAM J. Matrix Anal. Appl. 19 (2) (1998) 487–498.

[2] N.E. Barabanov, Lyapunov indicators of discrete inclusions, part I, II and III, Translation from Avtomat. i Telemekh. 2 (1988) 40–46., 3 (1988) 24–29 and 5 (1988) 17–24

[3] M.A. Berger, Y. Wang, Bounded semigroups of matrices, J. Linear Algebra Appl. 166 (1992) 21–27.

[4] V.D. Blondel, Y. Nesterov, Computationally efficient approximations of the joint spectral radius, Sixteenth International Symposium on Mathematical Theory of Networks and Systems (MTNS 2004), Leuven, Belgium, July 2004.

[5] I. Daubechies, J.C. Lagarias, Sets of matrices all infinite products of which converge, J. Linear Algebra Appl. 161 (1992) 227–263.

[6] W.P. Dayawansa, C.F. Martin, A converse lyapunov theorem for a class of dynamical systems which undergo switching, IEEE Trans. Automat. Control 44 (4) (1999) 751–760.

[7] G. Gripenberg, Computing the joint spectral radius, Linear Algebra Appl. 234 (1996) 43–60.

[8] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1993.

[9] Ralph Howard, The John Ellipsoid Theorem, University of South Carolina.

[10] F. John, Extremum problems with inequalities as subsidiary conditions, in: Studies and Essays Presented to R. Courant on his 60th birthday, January 8, 1948, Interscience Publishers, Inc., New York, NY, 1948, pp. 187–204.
[11] V.S. Kozyakin, Algebraic unsolvability of problem of absolute stability of desynchronized systems, Automat. Remote Control 51 (1990) 754–759.
[12] J.C. Lagarias, Y. Wang, The finiteness conjecture for the generalized spectral radius of a set of matrices, Linear Algebra Appl. 214 (1995) 17–42.
[13] D. Liberzon, J.P. Hespanha, A.S. Morse, Stability of switched systems: a Lie-algebraic condition, Syst. Control Lett. 37 (3) (1999) 117–122.
[14] M. Maesumi, An efficient lower bound for the generalized spectral radius, Linear Algebra Appl. 240 (1996) 1–7.
[15] Y. Nesterov, A. Nemirovsky, Interior Point Polynomial Algorithms in Convex Programming, Society for Industrial and Applied Mathematics, Philadelphia, 1994 (SIAM Studies in Applied Mathematics).
[16] H. Samelson, Notes on Lie Algebras, Van Nostrand Reinhold Co, New York, 1969.
[17] G. Strang, G.-C. Rota, A note on the joint spectral radius, Proc. Netherlands Acad. 22 (1960) 379–381.
[18] J. Theys, Note on upper and lower bounds for the joint spectral radius, Technical Report 2003.88, CESAME UCL (internal work, freely available on request).
[19] J.N. Tsitsiklis, in: T.M. Cover, B. Gopinath (Eds.), The Stability of the Products of a Finite Set of Matrices Open Problems in Communication and Computation, Springer-Verlag, New York, 1987, pp. 161–163.
[20] J. Tsitsiklis, V. Blondel, The Lyapunov exponent and joint spectral radius of pairs of matrices are hard—when not impossible—to compute and to approximate, Math. Control Signals Systems 10 (1997) 31–40. (Correction in 10 (1997) 381).
[21] F. Wirth, The generalized spectral radius and extremal norms, Linear Algebra Appl. 342 (2002) 17–40.